Quantum theory cannot violate a causal inequality

Tom Purves* and Anthony J. Short†
H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, U.K.

Within quantum theory, we can create superpositions of different causal orders of events, and observe interference between them. This raises the question of whether quantum theory can produce results that would be impossible to replicate with any classical causal model, thereby violating a causal inequality. This would be a temporal analogue of Bell inequality violation, which proves that no local hidden variable model can replicate quantum results. However, unlike the case of non-locality, we show that quantum experiments *can* be simulated by a classical causal model, and therefore cannot violate a causal inequality.

Introduction.– A fascinating aspect of quantum theory that has been investigated recently is the possibility for the causal order of events to be placed into superposition [1–5], leading to 'causal indefiniteness' about the order with which events have taken place. This phenomenon has been tested experimentally [6–8], and can be exploited to gain advantages within quantum theory. For example, setups based on the quantum switch [1] can help to determine whether unknown unitaries commute or anticommute [2]. An interesting question is whether quantum theory can generate results which could not be simulated by any classical causal model. Such results would violate a Causal Inequality [15–18]. These are the temporal analogues of Bell Inequalities [19], and the violation of such an inequality in nature would call into question the elementary properties that scientists regularly invoke when talking about cause and effect relationships.

In this paper we focus on the relationship between the type of causal indefiniteness present in quantum theory and the type needed to violate causal inequalities. We show that despite allowing causally indefinite processes, the correlations generated by quantum theory can be simulated by a classical causal model. This means that quantum theory cannot violate causal inequalities, and hence cannot yield an advantage over classical causal processes for tasks defined in a theory-independent way (such as 'guess your neighbour's input' [17, 20]). Previous works in this direction have shown that particular switch-type scenarios cannot violate causal inequalities [21], and that causal order cannot be placed in a pure superposition [13, 22]. It has also been shown that causal inequality violations are possible when we condition on measurement outcomes of one party [23]. However, our results imply that such violations are not possible for general quantum setups without conditioning.

Indefinite causal structure is often studied via process matrices [15], which assume that local laboratories obey standard quantum theory, but allow any connections between them consistent with this. This may include processes which are not achievable in standard quantum theory, or in nature more generally. Here we focus on what is possible in standard quantum theory, using quantum control of different parties' operations to generate superpositions of causal order, in a similar way to [24, 25]. As process matrices can yield causal inequality violation, a corollary of our result is that all process matrices cannot be implemented in standard quantum theory.

Results.– Before considering quantum processes, we first define causal processes, which are those which could be realised classically by a set of parties in separate laboratories passing systems between them [5, 26].

First consider two parties, Alice and Bob, with measurement settings x and y and measurement results a and b respectively. During the experiment, depicted in figure 1, each party sees a system enter their laboratory exactly once, performs a measurement on it with their corresponding measurement setting (which may also modify the system), and records their result. They then pass the system out of their laboratory. Apart from the systems entering and leaving their laboratories, the two parties cannot communicate with each other, but the system leaving one laboratory may be later sent into the other. Alice and Bob's joint measurement results can be described by a conditional probability distribution $p(ab|xy)$. However, not all such probability distributions can be achieved by a causal process.

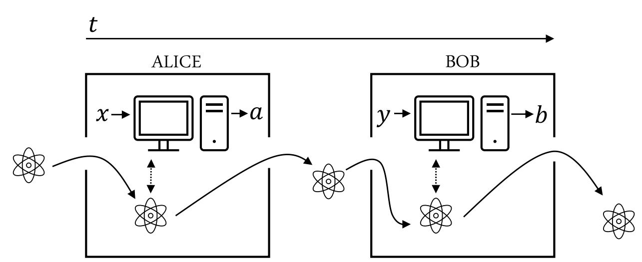

FIG. 1. An example of a causal process in which Alice goes before Bob. Note that the system which is passed from Alice's to Bob's laboratory could encode information about a and x.

The most general causal process in this case would be to first choose randomly whether Alice or Bob would go first (with probabilities $p(\text{Alice first})$ or $p(\text{Bob first})$). If Alice goes first, then her measurement result can depend on her measurement setting but not on Bob's, who hasn't acted yet, so is given by $p(a|x)$. She can then encode her

* tom.purves@bristol.ac.uk
† tony.short@bristol.ac.uk

measurement setting and result in the system and pass it out of her laboratory. This system then enters Bob's laboratory, where his measurement result can depend on all of the other variables, given by $p(b|a,x,y)$. Considering the other causal order in which Bob goes first in the same way, we obtain [17]

$$p^{\text{causal}}(ab|xy) = p(\text{Alice first})p(a|x)p(b|a,x,y) + p(\text{Bob first})p(b|y)p(a|b,x,y) \quad (1)$$

For the multiparty generalisation [18, 27], observe that the above causal probability contains two types of terms. The first, such as p(Alice first), determines the order in which the parties act, and the second, such as $p(a|x)$ or $p(b|a,x,y)$, determines the outcome probabilities of their measurements, constrained by their causal order. We now extend these 'who is next?' and 'what did they see?' type probabilities to an arbitrary number of parties. We use l_k to denote the kth party that receives the system (or equivalently, the kth laboratory the system enters), and denote the probability for this to occur by $p_k(l_k|H_{k-1})$. The conditional on H_{k-1} represents the history (including all previous parties that have measured, and their inputs and outputs) for it should be permitted for parties in the causal past of l_k to affect who is the next party to act. As a simple example of this, consider a tripartite experiment, with Alice, Bob and Charlie participating. If Charlie comes first, the system could be passed to Alice or Bob next, based on the outcome of his measurement. Here, $p_2(\text{Alice next}|\text{Charlie got outcome} = 1)$ may not be equal to $p_2(\text{Alice next}|\text{Charlie got outcome} = 0)$. Scenarios of this form this are what $p_k(l_k|H_{k-1})$ accounts for. The probability for l_k to obtain given results may also depend on this history (but, importantly, not on the causal future), and of course on the measurement setting, denoted x_{l_k}. We write this probability as $p_k(a_{l_k}|H_{k-1}, x_{l_k})$. A causal model is then the summation over all available parties at all stages of the measurement procedure, under the assumption that each party only acts once in the entire procedure.

Definition 1 *A causal probabilistic model can be written as*

$$p^{\text{causal}}(\vec{a}|\vec{x}) = \sum_{l_1 \notin \mathcal{L}_0} ... \sum_{l_N \notin \mathcal{L}_N - 1} p_1(l_1|H_0) p_1(a_{l_1}|H_0, x_{l_1})... \\ ...p_N(l_N|H_{N-1}) p_N(a_{l_N}|H_{N-1}, x_{l_N}) \quad (2)$$

where the $p_k(l_k|H_{k-1})$ terms represent probabilities for party l_k to act at stage k of the causal order, and $p_k(a_{l_k}|H_{k-1}, x_{l_k})$ terms represent probabilities for party l_k, who has acted at stage k of the causal order to obtain measurement result a_{l_k}. Both of the above probabilities are conditional on a history, H_{k-1}, which contains all of the information about previous inputs, outputs and party order. In particular, the history $H_k = (h_1, ..., h_k)$ is the ordered list of triples $h_i = (l_i, a_{l_i}, x_{l_i})$. The summations are performed over all possible next parties, excluding parties who have already acted, which are stored in the unordered sets $\mathcal{L}_k = \{l_1, ..., l_k\}$. To emphasise the symmetry between the terms we include H_0 and $\mathcal{L}_0$, which are defined as empty sets, as no parties have acted at that point.

This definition leads to a convex polytope of causal probability distributions $p^{\text{causal}}(\vec{a}|\vec{x})$. Note that although the notation differs, this generates the same set of probabilities as was previously defined in [18, 27]. Linear constraints on these probabilities which are satisfied by all $p^{\text{causal}}(\vec{a}|\vec{x})$ but which could be violated by some arbitrary probability distribution $p(\vec{a}|\vec{x})$ are known as 'causal inequalities', and are a temporal analogue of the Bell inequalities which have been widely studied in the context of quantum non-locality. By definition, any $p^{causal}(a,b|x,y)$ cannot violate a causal inequality. A violation of a causal inequality, by observation in experiment or by calculation in theory, proves that those experimental results or predictions do not have a causal explanation of the type defined above.

Quantum Processes.– A general representation of quantum theory is provided by the quantum circuit model. However, if we construct a circuit with the parties' actions at fixed locations, then there is no causal indefiniteness and a causal inequality cannot be violated [35]. Even to capture all classical causal processes, we need to be able to alter when different parties act. This can be achieved in the circuit model by representing the parties' actions by controlled quantum gates. Such gates could be constructed within standard quantum theory (e.g. by sending a system into the lab when the control is in the appropriate state and not otherwise), and are effectively what has been used in experiments probing quantum causality [7]. Quantum circuits involving controlled lab gates appear sufficient to represent any processes achievable within standard quantum theory.

For simplicity, we consider a setup involving a single quantum control which can trigger any of the labs. However this is equivalent to considering any quantum circuit which can be constructed from any number of individual controlled lab operations and other unitary gates (see the supplementary information for more details).

The key idea is to consider N parties, each of whom will interact with a quantum system exactly once, but in an order that is controlled coherently via the quantum control. We allow arbitrary unitary transformations of the system and control between each party's action, so that the ordering of later parties can be modified by earlier actions.

To allow the maximum possible interference, and avoid 'collapses' which would prevent interference between different causal orders, we model each party's measurement as a unitary interaction between the system and a local measurement register. This corresponds to the case in which there is no record in the measuring device of the time at which the measurement was performed. At the end of the experiment, all parties read off their measurement results from their local measurement registers

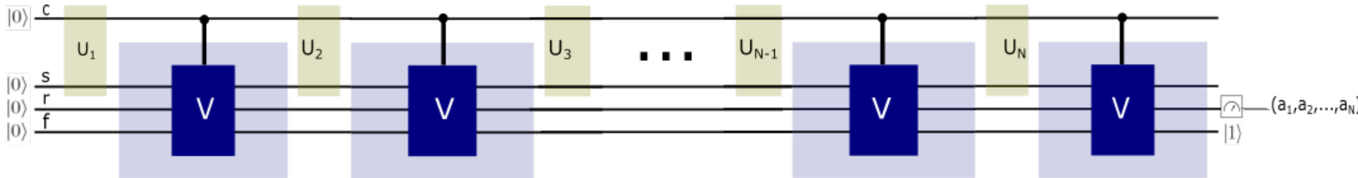

FIG. 2. Illustration of the quantum protocol. The system interacts with the different parties via a sequence of controlled entangling unitaries. Quantum control of causal order is achieved by a series of unitaries U_n on the control and system wires.

(which can be modelled by a standard projective measurement).

Each party also has a 'flag' which keeps track of how many times they have interacted with the system. At the end of the protocol we require that each party has interacted with the system exactly once.

Formally, the Hilbert space can be decomposed into the following components

- An arbitrary quantum system $\mathcal{H}_s$, which is passed between parties.
- A quantum control $\mathcal{H}_c$ which has dimension $N+1$. The basis states $|1\rangle \dots |N\rangle$ denote which party will measure next, while $|0\rangle$ is treated as a 'do nothing' command. By considering superpositions of these basis states, we can superpose different causal orders.
- A result register $\mathcal{H}_{r_i}$ for each of the N parties. The different results are represented by orthonormal basis states $|a_i\rangle$ with $a_i \in \mathcal{A}_i$, leading to the result register having dimension $|\mathcal{A}_i|$. We choose one of these states as a starting state for the results register and denote it by $|0\rangle_{r_i}$.
- A 'flag' $\mathcal{H}_{f_i}$ for each of the N parties, indicating how many times they have interacted with the system. For simplicity, we take each of these to be infinite dimensional, with basis states labelled by the integers. When the party interacts with the system the value of the flag is unitarily raised by the operator $\Gamma = \sum_n |n+1\rangle_{f_i} \langle n|_{f_i}$. Each flag starts in the $|0\rangle_{f_i}$ state, and at the end of the protocol, we require them all to be in the $|1\rangle_{f_i}$ state.

Note that we do not include separate local quantum ancillas for the parties, as these can always be incorporated in $\mathcal{H}_s$. We denote the combined result and flag spaces by $\mathcal{H}_r = \bigotimes \mathcal{H}_{r_i}$ and $\mathcal{H}_f = \bigotimes \mathcal{H}_{f_i}$ respectively.

We consider quantum protocols as follows. Firstly, the initial state

$$|0\rangle = |0\rangle_s |0\rangle_c |0\rangle_r^N |0\rangle_f^N \in \mathcal{H}_s \otimes \mathcal{H}_c \otimes \mathcal{H}_r \otimes \mathcal{H}_f. \quad (3)$$

is prepared, and each party l either chooses or is distributed their individual classical measurement setting x_l.

The protocol then consists of T time-steps, each of which is composed of two operations. Firstly, an arbitrary unitary transformation U_t is applied to the system and control, which can depend on the time t. Secondly, a fixed controlled lab-activation unitary V is applied, which activates whichever party is specified by the control. This is given by

$$V = |0\rangle\langle 0|_c \otimes I + \sum_{l=1}^{N} |l\rangle\langle l|_c \otimes V_{s,r_l}(x_l) \otimes \Gamma_{f_l} \otimes I \quad (4)$$

where the identities are over all remaining subsystems. $V_{s,r_l}(x_l)$ is a unitary which implements the measurement of party l on the system specified by the measurement setting x_l, and stores the result in the register r_l. For example, two different values of x_l could correspond to party l measuring the system in either the computational or the Fourier basis. Note that by incorporating ancillas within the system, any local quantum measurement (i.e a POVM) is realisable within this paradigm. Ancillas can also be used to generate arbitrary mixed states if required (via purification).

The unitary operator Γ_{f_l} raises the flag system of the party making their measurement. At the end of the protocol, we require that the flags are in the state $|1\rangle_f^N$ (i.e. that each party has measured the system once). This places constraints on the possible protocols which can be constructed. Note that each party does not have access to an operation which resets the flag, aside from the initialisation operation at the start of the protocol. They therefore always 'remember' if they have made a measurement or not. Also, we do not allow circuits involving the controlled inverse of a party's action (which would lower their flag and erase their memory), as this would enlarge the set of causal possibilities even classically.

The total unitary for the protocol is given by

$$\mathcal{U} = VU_T VU_{T-1} \dots VU_1. \quad (5)$$

At the end of the protocol, each party performs a projective measurement on their results register to obtain their final result [36]. The output probability distribution of the quantum protocol is therefore given by

$$p^{\text{quantum}}(\vec{a}|\vec{x}) = |(|\vec{a}\rangle\langle\vec{a}|_r \otimes I)\mathcal{U}|0\rangle|^2. \quad (6)$$

The full protocol is illustrated as a quantum circuit in figure 2.

The main result of this paper is that any probability distribution which can be generated within quantum theory, as described above, can also be obtained via a classical causal process.

Theorem 1 *Any quantum probability distribution $p^{\text{quantum}}(\vec{a}|\vec{x})$ can be exactly replicated by a classically causal process $p^{\text{causal}}(\vec{a}|\vec{x})$. Hence quantum theory cannot violate a causal inequality.*

In particular, we now show how to construct an explicit classical causal process which replicates the results of any quantum protocol, together with a sketch of the proof of Theorem 1. The full proof of the theorem can be found in the supplementary information.

We first define notation for describing states at each stage of the quantum protocol, and then show how to use these to construct the probabilities in the corresponding classical model.

Definition 2 *The (un-normalised) state with a History H_{k-1}, at a time t, with the control set to trigger the action of party l_k is given by*

$$|\psi_{(l_k,t,H_{k-1})}\rangle = (|l_k\rangle\langle l_k|_c \otimes \pi_{rf}^{H_{k-1}} \otimes I_s)U_t V U_{t-1}...VU_1|0\rangle\,. \quad (7)$$

The projector onto the result and flag spaces is given by $\pi_{rf}^{H_{k-1}} = \bigotimes_{i=1}^{N}\left(\pi_{r_i f_i}^{H_{k-1}}\right)$, where

$$\pi_{r_i f_i}^{H_{k-1}} = \begin{cases} |a_i\rangle\langle a_i|_{r_i} \otimes |1\rangle\langle 1|_{f_i} & \text{if } (i, a_i, x_i) \in H_{k-1}, \\ I_{r_i} \otimes |0\rangle\langle 0|_{f_i} & \text{otherwise}\,. \end{cases} \quad (8)$$

This notation describes states which are about to be measured by the parties (i.e., a V type operator is about to act on them). We also set up some notation for states which have just been measured, in a similar fashion.

Definition 3 *The (un-normalised) state with a History H_k, at a time t, in which party l_k has just acted is given by*

$$|\phi_{(l_k,t,H_k)}\rangle = (|a_{l_k}\rangle\langle a_{l_k}|_{r_k} \otimes I)V|\psi_{(l_k,t,H_{k-1})}\rangle\,. \quad (9)$$

With these definitions, we can associate the states in this quantum process with the probabilities in our classical causal model.

Definition 4 *The probability for party l_k to act next, given a history H_{k-1} is given by:*

$$p_k(l_k|H_{k-1}) = \frac{\sum_{t_k=1}^{T} |\,|\psi_{(l_k,t_k,H_{k-1})}\rangle\,|^2}{\sum_{l'_k \notin \mathcal{L}_{k-1}} \sum_{t'_k=1}^{T} |\,|\psi_{(l'_k,t'_k,H_{k-1})}\rangle\,|^2}. \quad (10)$$

We have summed over time [36], because it is possible within the quantum paradigm to conduct the k^{th} measurement at different times according to a background clock (which we note the labs have no access to). Note that states at different times combine incoherently, but different sequences leading to the same set of historical measurement results combine coherently inside $|\psi_{(l_k,t_k,H_{k-1})}\rangle$.

The form of equation (10) makes it a valid probability distribution, as it is non-negative, and obeys the correct normalisation that $\sum_{l_k \notin \mathcal{L}_{k-1}} p_k(l_k|H_{k-1}) = 1$. Also note that it depends on only those input variables x_i which appear in the history H_{k-1}.

Next, we specify similar probabilities for seeing measurement results based on a given history.

Definition 5 *The probability for party l_k to obtain the measurement result a_{l_k}, given a history H_{k-1}, and an input variable x_{l_k} is given by:*

$$p_k(a_{l_k}|H_{k-1}, x_{l_k}) = \frac{\sum_{t_k=1}^{T} |\,|\phi_{(l_k,t_k,H_k)}\rangle\,|^2}{\sum_{a'_{l_k} \in \mathcal{A}_{l_k}} \sum_{t'_k=1}^{T} |\,|\phi_{(l_k,t'_k,H'_k)}\rangle\,|^2}, \quad (11)$$

where $H_k = (H_{k-1}, (l_k, a_{l_k}, x_{l_k}))$ and $H'_k = (H_{k-1}, (l_k, a'_{l_k}, x_{l_k}))$

This is again a valid probability distribution, since $\sum_{a_{l_k} \in \mathcal{A}_{l_k}} p_k(a_{l_k}|H_{k-1}, x_{l_k}) = 1$. In the numerator, we have simply taken sum of the modulus squared of all of the states which have the correct historical results, the control in the correct state, and the results register containing the result we want to calculate the probability for.

To prove Theorem 1, We begin by inserting $p_k(l_k|H_{k-1})$ (from (10)) and $p_k(a_{l_k}|H_{k-1}, x_{l_k})$ (from (11)) into the definition of a causal model (1). We are then able to straightforwardly cancel the numerator of the 'who is next?' type probabilities with the denominator of the 'what did they see?' probabilities for the probabilities evaluated *at the same stage of the causal order*. Next, we show that a sum over the last party to measure in the numerator at one stage of the causal order, cancels with the denominator *at the next stage of the causal order*. We then make the observation that for the first stage of the causal order, the denominator of $p_1(l_1|H_0)$ is equal to one (which corresponds to the fact that someone must measure first in the quantum circuit). Finally, we note that the numerator of the final term, summed over all parties, represents exactly the probabilities $p^{\text{quantum}}(\vec{a}|\vec{x})$ arising from the quantum protocol. This allows us to simulate the results of the quantum protocol via the classically causal model given in (1). Given that it can be replicated by a causal model, it follows that quantum theory cannot violate a causal inequality.

In the supplementary information, we give an example of how these results can be applied in practice, based on the quantum switch [1]. This involves the causal order of two parties becoming entangled with the control. A third party then performs a measurement which leads to interference between the two causal orders. It has already been shown that this simple setup cannot be used to violate a causal inequality [15, 21]. However, it is instructive to see how it fits into our framework. Despite the quantum setup including interference, our results give an explicit classical causal process which generates the same behaviour (i.e. the same $p(a, b, c|x, y, z)$).

Conclusions.– By using a quantum control to determine when different parties measure, and treating these measurements as coherent unitary processes, quantum theory allows us to generate superpositions of causal orders and to observe interference between them. At the level of the theory, such processes do not arise from a single causal order, or even a mixture of orders. However, we have shown that the probabilities $p(\vec{a}|\vec{x})$ generated by any quantum protocol *can* be simulated by a classical causal process. This means that quantum theory cannot violate a causal inequality, and thus one could not convince a sceptic that nature deviates from classical notions of causality.

This is in sharp distinction to non-locality, where not only does the theory appear non-local (e.g. via entangled states) but we can also prove that some quantum probabilities cannot be replicated by any local hidden variable model. By violating a Bell inequality we can therefore prove non-locality experimentally.

Although our framework is very general, one key requirement is that each party interacts with the system once (which leads to a requirement on the final flag state). This is the normal setup for causal inequalities, and allows us to assign a single input and output to each party, and to represent the experimental results via $p(\vec{a}|\vec{x})$. However, it would be interesting to lift this assumption in future research. For example, could we obtain a violation of causality if parties are allowed to measure twice, or a variable number of times, or to forget they have measured? We also have a technical assumption that the protocol takes finite time (i.e. that it terminates after a finite number of steps). This seems physically reasonable, but it might be interesting to investigate lifting this assumption, as well as to consider extending the results to continuous time. Finally, it would be interesting to consider a network structure in the causal scenario, in the non-local case this is known to generate non-linear Bell inequalities, and sets of non convex probability polytopes [33]. Investigation of causal indefiniteness and causal inequalities in these type of scenarios might prove of general interest.

Finally, our framework assumes standard quantum theory. If the theory changes significantly to incorporate quantum gravity we might expect new possibilities for causal inequality violation, although not necessarily [4] (note that even classical general relativity allows for the existence of closed time-like curves, which appear to violate the simple classical causal models we have considered here [14, 28, 29]). We hope that the framework and tools developed here will prove helpful in discussing these interesting issues, and in highlighting differences from the standard case.

T.Purves acknowledges support from the EPSRC.

Note added - Independently obtained related results using the process matrix formalism [34] appeared on the ArXiv on the same day as this paper.

[1] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Physical Review A*, 88(2), 2013.

[2] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Phys. Rev. A*, 86:040301(R), Oct 2012.

[3] Cyril Branciard. Witnesses of causal nonseparability: an introduction and a few case studies. *Scientific Reports*, 6(1):26018, May 2016.

[4] Magdalena Zych, Fabio Costa, Igor Pikovski, and Časlav Brukner. Bell's theorem for temporal order. *Nature Communications*, 10(1):3772, Aug 2019.

[5] Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. Quantum causal models, 2020. arXiv:1906.10726.

[6] Lorenzo M. Procopio, Amir Moqanaki, Mateus Araújo, Fabio Costa, Irati Alonso Calafell, Emma G. Dowd, Deny R. Hamel, Lee A. Rozema, Časlav Brukner, and Philip Walther. Experimental superposition of orders of quantum gates. *Nature Communications*, 6(1), Aug 2015.

[7] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White. Indefinite causal order in a quantum switch. *Phys. Rev. Lett.*, 121:090503, 2018.

[8] K. Goswami and J. Romero. Experiments on quantum causality. *AVS Quantum Science*, 2(3):037101, Oct 2020.

[9] Aram W. Harrow, Avinatan Hassidim, Debbie W. Leung, and John Watrous. Adaptive versus nonadaptive strategies for quantum channel discrimination. *Phys. Rev. A*, 81:032339, Mar 2010.

[10] Daniel Ebler, Sina Salek, and Giulio Chiribella. Enhanced communication with the assistance of indefinite causal order. *Phys. Rev. Lett.*, 120:120502, 2018.

[11] Guérin, Philippe Allard and Rubino, Giulia and Brukner, Časlav Communication through quantum-controlled noise *Phys. Rev. A*, 99.062317, Jun 2019.

[12] Rubino, Giulia and Rozema, Lee A. and Ebler, Daniel and Kristjánsson, Hlér and Salek, Sina and Allard Guérin, Philippe and Abbott, Alastair A. and Branciard, Cyril and Brukner, Časlav and Chiribella, Giulio and et al. Experimental quantum communication enhancement by superposing trajectories *Physical Review Research*, 3.013093, Jan 2021.

[13] Yokojima, Wataru and Quintino, Marco Túlio and Soeda, Akihito and Murao, Mio Consequences of preserving reversibility in quantum superchannels *Quantum*, q-2021-04-26-441, Apr 2021.

[14] Araújo, Mateus and Guérin, Philippe Allard and Baumeler, Ämin Quantum computation with indefinite causal structures *Physical Review A*, 96.052315, Nov 2017.

[15] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3(1):1092, 2012.

[16] Časlav Brukner. Quantum causality. *Nature Physics*, 10(4):259–263, 2014.

[17] Cyril Branciard, Mateus Araújo, Adrien Feix, Fabio Costa, and Časlav Brukner. The simplest causal inequalities and their violation. *New Journal of Physics*, 18(1):013008, 2015.

[18] Alastair A. Abbott, Christina Giarmatzi, Fabio Costa, and Cyril Branciard. Multipartite causal correlations: Polytopes and inequalities. *Physical Review A*, 94(3), Sep 2016.

[19] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, 1964.

[20] Mafalda L. Almeida, Jean-Daniel Bancal, Nicolas Brunner, Antonio Acín, Nicolas Gisin, and Stefano Pironio. Guess your neighbor's input: A multipartite nonlocal game with no quantum advantage. *Physical Review Letters*, 104(23), Jun 2010.

[21] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. Witnessing causal nonseparability. *New Journal of Physics*, 17(10):102001, Oct 2015.

[22] Fabio Costa. A no-go theorem for superpositions of causal orders, 2020. arXiv:2008.06205.

[23] Simon Milz, Dominic Jurkschat, Felix A. Pollock, and Kavan Modi. Quantum chicken-egg dilemmas: Delayed-choice causal order and the reality of causal nonseparability, 2020. arXiv:2008.07876.

[24] Timoteo Colnaghi, Giacomo Mauro D'Ariano, Stefano Facchini, and Paolo Perinotti. Quantum computation with programmable connections between gates. *Physics Letters A*, 376(45):2940 – 2943, 2012.

[25] Mateus Araújo, Fabio Costa, and Časlav Brukner. Computational advantage from quantum-controlled ordering of gates. *Phys. Rev. Lett.*, 113:250402, Dec 2014.

[26] Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009.

[27] Ognyan Oreshkov and Christina Giarmatzi. Causal and causally separable processes. *New Journal of Physics*, 18(9):093020, 2016.

[28] Francisco S. N. Lobo. Closed timelike curves and causality violation, 2010. arXiv:1008.1127.

[29] Jonathan Barrett, Robin Lorenz, and Ognyan Oreshkov. Cyclic quantum causal models, 2020. arXiv:2002.12157.

[30] Hugh Everett. "relative state" formulation of quantum mechanics. *Rev. Mod. Phys.*, 29:454–462, Jul 1957.

[31] Note that from a many-worlds perspective [30] such an additional step would not be necessary. However, we include it here to maintain connection with standard quantum theory and give an explicit formula for the outcome probabilities.

[32] Here, we summed from $t = 1$ to T, but we could have equivalently chosen $t = k$ to $T - (N - k)$, and only went with the former to make the subsequent notation easier to read, at the cost of including 0 amplitude states in our probability definition.

[33] Nicolas Gisin, Jean-Daniel Bancal, Yu Cai, Patrick Remy, Armin Tavakoli, Emmanuel Zambrini Cruzeiro, Sandu Popescu, Nicolas Brunner. Constraints on nonlocality in networks from no-signaling and independence. *Nature Communications*, 11(1):2378, 2020

[34] Julian Wechs and Hippolyte Dourdent and Alastair A. Abbott and Cyril Branciard. Quantum circuits with classical versus quantum control of causal order arXiv:2101.08796.

[35] We could space out the circuit such that there is at most one lab at each time-step, and then pass the full quantum state between the labs as a classical hidden variable which would allow us to recover the same correlations in a causal way.

[36] Note that from a many-worlds perspective [30] such an additional step would not be necessary. However, we include it here to maintain connection with standard quantum theory and give an explicit formula for the outcome probabilities